WHAT IS CLAIMED IS:

1.   A data classification method of classifying a group of data into a plurality of sets in accordance with

5   data values, comprising:

dividing said group of data into a first number of sets having no common elements; and

calculating a first total degree of randomness which is a sum of degrees of randomness of said data

10   values in said respective sets of said first number of sets,

wherein data division to said first number of sets and calculation of said first total degree of randomness are repeated while a form of data division to said first

15   number of sets is changed, and said group of data is classified into data belonging to the respective classification sets of said first number of classification sets in which said first total degree of randomness is minimized.

20

2.   The method according to claim 1, wherein data division to said first number of sets is performed for data to be classified in numerical order of data values.

25   3.   The method according to claim 1, wherein said calculating the sum of degrees of randomness in the respective sets of said first number of sets comprises:

estimating a probability distribution of data

values in each of said sets on the basis of said data

values of said data belonging to each of said sets;

obtaining an entropy of each of said estimated

probability distributions of data values; and

5   weighting said entropy of each of said probability

distributions in accordance with the number of data

belonging to a corresponding one of said sets.

4.   The method according to claim 3, wherein said

10   first probability distribution is a normal distribution.

5.   The method according to claim 1, further

comprising:

dividing data belonging to a specific

15   classification set in said first number of classification

sets into a second number of sets having no common

elements; and

calculating a second total degree of randomness

which is a sum of degrees of randomness of data values in

20   the respective sets of said second number of sets,

wherein data division to said second number of sets

and calculation of said second total degree of randomness

are repeated while a form of data division to said second

number of sets is changed, and said data belonging to

25   said specific classification set are further classified

into data belonging to the respective classification sets

of said second number of classification sets in which

said second total degree of randomness is minimized.

6. The method according to claim 5, wherein data division to said second number of sets is performed for data to be classified in numerical order of data values.

7. The method according to claim 5, wherein said calculating the sum of degrees of randomness in the respective sets of said second number of sets comprises:

estimating a probability distribution of data values in each of the sets on the basis of said data values of said data belonging to each of said sets;

obtaining an entropy of each of the estimated probability distributions of data values; and

weighting said entropy of each of said probability distributions in accordance with the number of data belonging to a corresponding one of said sets.

8. The method according to claim 7, wherein said first probability distribution is a normal distribution.

9. A data classification apparatus for classifying a group of data into a plurality of sets in accordance with data values, comprising:

a first data dividing unit which divides said group of data into a first number of sets having no common elements; and

a first degree-of-randomness calculation unit which calculates degrees of randomness of data values in the

respective sets divided by said first data dividing unit, and calculates a sum of the degrees of randomness; and

a first classification unit which classifies said group of data into said data belonging to the respective

5 classification sets of said first number of classification sets in which said sum of degrees of randomness calculated by said first degree-of-randomness calculation unit is minimum out of forms of data division by said first data dividing unit.

10

10. The apparatus according to claim 9, further comprising:

a second data dividing unit which divides data belonging to a specific classification set in the first

15 number of classification sets into a second number of sets having no common elements; and

a second degree-of-randomness calculation unit which calculates degrees of randomness of data values in the respective sets divided by said second data dividing

20 unit and calculates a sum of the degrees of randomness; and

a second classification unit which classifies said data of said specific classification set into said data belonging to the respective classification sets of said

25 second number of classification sets in which said sum of degrees of randomness calculated by said second degree-of-randomness calculation unit is minimum out of forms of data division by said second data dividing unit.

11. A signal processing method of processing a measurement signal obtained by measuring an object, comprising:

5   extracting signal levels at a plurality of feature points obtained from said measurement signal; and

setting said extracted signal levels as classification object data and classifying said signal levels at said group of feature points into a plurality

10   of sets by using the data classification method according to claim 1.

12. The method according to claim 11, wherein said feature point is at least one of a local maximum point

15   and a local minimum point of said measurement signal.

13. The method according to claim 11, wherein said feature point is a point of inflection of said measurement signal.

20

14. A signal processing apparatus for processing a measurement signal obtained by measuring an object, comprising:

a measurement unit which measures said object and

25   acquires a measurement signal;

an extraction unit which extracts signal levels at a plurality of feature points obtained from said measurement signal; and

the data classification apparatus according to
claim 9, which sets said extracted signal levels as
classification object data.

5   15.   A position detection method of detecting a
position of a mark formed on an object, comprising:
    acquiring an image pick-up signal by picking up an
image of said mark;
    processing said image pick-up signal as a
10 measurement signal by said signal processing method
according to claim 11; and
    calculating said position of said mark on the basis
of a signal processing result obtained in said signal
processing.

15

    16.   The method according to claim 15, wherein
    in data classification in said signal processing,
the number of data which should belong to each
classification set after said data classification is
20 known in advance, and
    in said position calculation, the number of data
which should belong to each classification set is
compared with the number of data in each of said
classification sets classified in said signal processing
25 to evaluate validity of the classification in said signal
processing, and said position is calculated on the basis
of said data belonging to said classification set
evaluated to be valid.

17. A position detection apparatus for detecting a position of a mark formed on an object, comprising:

an image pick-up unit which acquires an image pick-up signal by picking up an image of said mark;

5 the signal processing apparatus according to claim 14, which performs signal processing for said image pick-up signal as a measurement signal; and

a position calculation unit which calculates said position of said mark on the basis of a signal processing
10 result obtained by said signal processing apparatus.

18. An exposure method of transferring a predetermined pattern onto a divided area on a substrate,
15 comprising:

detecting a position of a position detection mark formed on said substrate by the position detection method according to claim 15, obtaining a predetermined number of parameters associated with a position of said divided
20 area, and calculating arrangement information of said divided area on said substrate; and

transferring said pattern onto said divided area while performing position control on said substrate on the basis of said arrangement information of said divided
25 area obtained in said arrangement calculation.

19. An exposure apparatus for transferring a predetermined pattern onto a divided area on a substrate,

comprising:

a substrate stage on which said substrate is mounted; and

the position detection apparatus according to claim

5 17, which detects a position of said mark on said substrate.

20. A data classification method of classifying a group of data into a plurality of sets in accordance with

10 data values, comprising:

classifying said group of data into a first number of sets in accordance with said data values; and

dividing said group of data again into a second number of sets which is smaller than said first number on

15 the basis of a characteristic of each of said first number of sets divided in data classification into said first number of sets.

21. The method according to claim 20, wherein data

20 classification into said second number of sets comprises:

specifying a first set, of said first number of sets, which meets a predetermined condition;

estimating a first boundary candidate for dividing said group of data excluding data included in said first

25 set by using a predetermined estimation technique;

estimating a second boundary candidate for dividing a data group, of said group of data, which is divided by said first boundary candidate and includes said first set

by using said predetermined estimation technique; and

dividing said group of data into said second number of sets on the basis of said second boundary candidate.

5   22.   The method according to claim 21, wherein said predetermined estimation technique comprises:

calculating a degree of randomness of data values in each set divided by said boundary candidate, and calculating a sum of said degrees of randomness; and

10   performing said degree-of-randomness calculation while changing a form of data division with said boundary candidate, and extracting a boundary candidate with which said sum of degrees of randomness obtained in said degree-of-randomness calculation is minimized.

15

23.   The method according to claim 21, wherein said predetermined estimation technique comprises:

obtaining a probability distribution in each set of said data group; and

20   extracting said boundary candidate on the basis of a point of intersection of said probability distributions of the respective sets.

24.   The method according to claim 21, wherein said

25 predetermined estimation technique comprises:

calculating an inter-class variance as a variance between sets divided by said boundary candidate; and

performing said intra-class variance calculation

while changing a form of data division with said boundary candidate, and extracting a boundary candidate with which the inter-class variance obtained in said inter-class variance calculation is maximized.

5

    25.   The method according to claim 21, wherein said predetermined condition is a condition that data exhibiting a value substantially equal to a predetermined value is extracted from said group of data.

10

    26.   The method according to claim 25, wherein

    said group of data is image pick-up data of the respective pixels obtained by picking up different image patterns within a predetermined image pick-up field; and

15   said predetermined value is image pick-up data of a pixel existing in an area corresponding to an image pick-up area for a predetermined image pattern.

    27.   The method according to claim 20, wherein said

20 dividing data into said second number of sets comprises:

    extracting a predetermined number of sets from the first number of sets on the basis of the number of data included in the respective sets of said first number of sets;

25   calculating an average data value by averaging data values respectively representing sets of said predetermined number of sets; and

    dividing said group of data into said second number

113

of sets on the basis of said average data value.

28.   The method according to claim 27, wherein in
said average data value calculation, a weighted average
5   of said data values is calculated by using a weight
corresponding to at least one of the number of data of
the respective sets of said predetermined number of sets
and a probability distribution of said predetermined
number of sets.

10   29.
28.   The method according to claim 20, wherein said
first number is not less than three, and said second
number is two.

30.
15   29.   The method according to claim 20, wherein said
group of data is luminance data of the respective pixels
obtained by picking up different image patterns within a
predetermined image pick-up field.

31.
20   30.   A data classification apparatus for classifying
a group of data into a plurality of sets in accordance
with data values, comprising:
a first data dividing unit which divides said group
of data into a first number of sets on the basis of said
25   data values; and
a second data dividing unit which divides said
group of data into a second number of sets smaller than
said first number again on the basis of a characteristic

of each of said first number of sets.

31. The method according to claim 30, wherein said first number is not less than three, and said second

5 number is two.

32. An image processing method of processing image data obtained by picking up an image in a predetermined image pick-up field, comprising:

10 setting luminance data, as a group of data, which is obtained by picking up an image pattern of an object and an image pattern of a background which exist in said predetermined image pick-up field; and

identifying a boundary between said object and said

15 background by classifying said luminance data by using the data classification method according to claim 29.

33. The method according to claim 32, wherein said object includes a substrate onto which a predetermined

20 pattern is transferred.

34. An image processing apparatus for processing image data obtained by picking up an image in a predetermined image pick-up field, wherein

25 luminance data, which is obtained by picking up an image pattern of an object and an image pattern of a background which exist in said predetermined image pick-up field is set as a group of data, and

a boundary between said object and said background
is identified by classifying said luminance data by using
the data classification apparatus according to claim 30.

34.

35. An exposure method of transferring a
5   predetermined pattern onto a substrate, comprising:
        specifying an outer shape of said substrate by
using the image processing method according to claim 33;
        controlling a rotational position of said substrate
10 on the basis of said specified outer shape of said
substrate;
        detecting a mark formed on said substrate after
said rotational position is controlled; and
        transferring said predetermined pattern onto said
15 substrate while positioning said substrate on the basis
of a mark detection result obtained in said mark
detection.

35.

36. An exposure apparatus for transferring a
20 predetermined pattern onto a substrate, comprising:
        an outer shape specifying unit including the image
processing apparatus according to claim 34, which
specifies an outer shape of said substrate;
        a rotational position control unit which controls a
25 rotational position of said substrate on the basis of
said outer shape of said substrate which is specified by
said image processing apparatus;
        a mark detection unit which detects a mark formed

on said substrate whose rotational position is controlled by said rotational position control unit; and

a positioning unit which positions said substrate on the basis of a mark detection result obtained by said
5   mark position detection unit,

wherein said predetermined pattern is transferred onto said substrate while said substrate is positioned by said positioning unit.

10   37.   A data classification method of classifying a group of data into a plurality of sets in accordance with data values, comprising:

estimating a first number of boundary candidates for dividing said group of data into a second number of
15   sets on the basis of said data values; and

extracting a third number of boundary candidates which is smaller than said first number and is used to divide said group of data into a fourth number of sets smaller than said second number, under a predetermined
20   extraction condition, on the basis of said first number of boundary candidates.

38.   The method according to claim 37, wherein said predetermined extraction condition includes a condition
25   that said third number of boundary candidates are extracted on the basis of a magnitude of a data value indicated by each of said first number of boundary candidates.

39. The method according to claim 38, wherein said predetermined extraction condition includes a condition that a boundary candidate with which said data value is

5   maximized is extracted.

40. The method according to claim 37, wherein

said group of data are arranged at positions in a predetermined direction, and

10   said predetermined extraction condition includes a condition that said fourth number of boundary candidates are extracted on the basis of the respective positions of said first number of boundary candidates.

41. The method according to claim 37, wherein

15   said group of data are differential data obtained by differentiating image pick-up data of the respective pixels obtained by picking up different image patterns in a predetermined image pick-up field in accordance with

20   positions of said pixels,

said data value is a differential value of said image pick-up data, and

said boundary candidate is a position of said pixel.

42. The method according to claim 37, wherein said

25   first number is not less than two, and said third number is one.

44,
4̶3̶. The method according to claim 37, wherein said group of data are luminance data of the respective pixels obtained by picking up different image patterns in a predetermined image pick-up field.

5

45
4̶4̶. A data classification apparatus for classifying a group of data into a plurality of sets in accordance with data values, comprising:

a first data dividing unit which estimates a first
10 number of boundary candidates for dividing said group of data into a second number of sets on the basis of said data values; and

a second data dividing unit which extracts a third number of boundary candidates which is smaller than said
15 first number and is used to divide said group of data into a fourth number of sets smaller than said second number, under a predetermined extraction condition, on the basis of said first number of boundary candidates.

46.
20 4̶5̶. The apparatus according to claim 44, wherein said group of data are differential data obtained by differentiating image pick-up data of the respective pixels obtained by picking up different image patterns in a predetermined image pick-up field in accordance with
25 positions of said pixels,

said data value is a differential value of said image pick-up data, and

said boundary candidate is a position of said pixel.

predetermined pattern onto a substrate, comprising:

specifying an outer shape of said substrate by using the image processing method according to claim 47;

controlling a rotational position of said substrate

5   on the basis of said specified outer shape of said

substrate;

detecting a mark formed on said substrate after said rotational position is controlled; and

transferring said predetermined pattern onto said

10   substrate while positioning said substrate on the basis

of a mark detection result obtained in said mark

detection.

50.   An exposure apparatus for transferring a

15   predetermined pattern onto a substrate, comprising:

an outer shape specifying unit including the image

processing apparatus according to claim 48, which

specifies an outer shape of said substrate;

a rotational position control unit which controls a

20   rotational position of said substrate on the basis of

said outer shape of said substrate which is specified by

said image processing apparatus;

a mark detection unit which detects a mark formed

on said substrate whose rotational position is controlled

25   by said rotational position control unit; and

a positioning unit which positions said substrate

on the basis of a mark detection result obtained by said

mark position detection unit,

wherein said predetermined pattern is transferred onto said substrate while said substrate is positioned by said positioning unit.

5   52.
51.   A recording medium on which a position detection control program executed by a position detection apparatus for detecting a position of a mark formed on an object is recorded, wherein

said position detection control program comprises:

10   allowing an image of said mark to be picked up and allowing an image pick-up signal to be acquired;

a signal processing control program using said image pick-up signal as a measurement signal, comprising

allowing signal levels at a plurality of feature
15   points obtained from said measurement signal to be extracted; and

said data classification control program using said extracted signal levels as a group of classification object data, comprising

20   allowing said group of data to be divided into a first number of sets having no common elements;

allowing a first total degree of randomness which is a sum of degrees of randomness of data values in the respective sets of said first number of sets to be
25   calculated; and

allowing said group of data to be divided into data belonging to the respective classification sets of said first number of classification sets in which said first

total degree of randomness is minimized, by repeating

data division to said first number and calculation of

said first total degree of randomness while changing a

mode of data division to said first number of sets; and

5   allowing a position of said mark to be calculated

on the basis of a processing result on said image pick-up

signal.

52.   The medium according to claim 51, wherein

10   in said data classifying, the number of data which

should belong to each classification set after said data

classification is known in advance, and

        the number of data which should belong to each

classification set is compared with the number of data in

15   each of said classified classification sets to evaluate

validity of said data classifying, and said position is

calculated on the basis of data belonging to said

classification set evaluated to be valid.

53.   A recording medium on which an image processing

20   control program executed by an image processing apparatus

for processing image data obtained by picking up an image

in a predetermined image pick-up field is recorded,

wherein

25   said image processing control program comprises:

        allowing luminance data, which is obtained by

picking up an image pattern of an object and an image

pattern of a background which exist in said predetermined

image pick-up field, to be set as a group of data;

a data classification control program which allows said luminance data to be classified, comprising:

allowing said group of data to be divided into a
5 first number of sets on the basis of said data values; and

allowing said group of data to be divided into a second number of sets smaller than said first number again on the basis of features of the respective first
10 number of sets; and

allowing a boundary between said object and said background to be identified.

55.
54. A recording medium on which an image processing
15 control program executed by an image processing apparatus for processing image data obtained by picking up an image in a predetermined image pick-up field is recorded, wherein

said image processing control program comprises:
20 allowing luminance data which is obtained by picking up an image pattern of an object and an image pattern of a background which exist in said predetermined image pick-up field to be set as a group of data;

a data classification control program which allows
25 said luminance data to be classified, comprising

allowing a first number of boundary candidates for dividing said group of data into a second number of sets to be estimated on the basis of said data values;

allowing a third number of boundary candidates which is smaller than said first number and is used to divide said group of data into a fourth number of sets smaller than said second number, under a predetermined

5 extraction condition, to be extracted on the basis of said first number of boundary candidates; and

allowing a boundary between said object and said background to be identified.

10 55. A device manufacturing method including a lithography process, wherein

exposure is performed by using the exposure method according to claim 18 in said lithography process.

15 56. A device manufacturing method including a lithography process, wherein

exposure is performed by using the exposure method according to claim 35 in said lithography process.

20 57. A device manufacturing method including a lithography process, wherein

exposure is performed by using the exposure method according to claim 49 in said lithography process.